AVERAGING ANALYSIS FOR DISCRETE TIME AND SAMPLED DATA ADAPTIVE SYSTEMS

by

Er-Wei Bai, Li-Chen Fu and S. Shankar Sastry

Memorandum No. UCB/ERL M86/67

2 September 1986

COVER PAGE

AVERAGING ANALYSIS FOR DISCRETE TIME AND SAMPLED DATA ADAPTIVE SYSTEMS

by

Er-Wei Bai, Li-Chen Fu and S. Shankar Sastry

Memorandum No. UCB/ERL M86/67

2 September 1986

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

AVERAGING ANALYSIS FOR DISCRETE TIME AND SAMPLED DATA ADAPTIVE SYSTEMS

by

Er-Wei Bai, Li-Chen Fu and S. Shankar Sastry

Memorandum No. UCB/ERL M86/67

2 September 1986

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Averaging Analysis for Discrete Time and Sampled Data Adaptive Systems

*Er-Wei Bai, Li-Chen Fu and S. Shankar Sastry**

Electronic Research Laboratory
Department of Electrical Engineering & Computer Science
University of Berkeley,
Berkeley, CA 94720

ABSTRACT

We extend our earlier continuous time averaging theorems to the nonlinear discrete time case. We use theorems for the study of the convergence analysis of discrete time adaptive identification and control systems. We also derive instability theorems and use them for the study of robust stability and instability of adaptive control schemes applied to sampled data systems. As a by product we also study the effects of sampling on unmodeled dynamics in continuous time systems.

September 2, 1986

Keywords: Averaging, Lyapunov Converse Theorem, Adaptive Identification, G.R.C. Adaptive Control Scheme, Parameter Convergence, Tuned System, Slow-Drift Instability, Sampled-Data System

1 Introduction

In this paper, which we consider partially a review paper, we develop averaging theory for more general discrete time systems. The results are an extension to the nonlinear discrete time case of the continuous time results presented by several authors[1,16,17,18,22-24,26-31] and most recently in our own work Fu et al and Bodson et al [6,9,10].

The novelty of this paper is that the application of these results is not limited to linear or linearized systems, but allows for the analysis of the full nonlinear systems. We apply these results to some discrete time adaptive identification and control schemes that have recently appeared in the literature, namely, projection type and least squares type identifiers, and the d-step ahead control law with the projection or least squares identifier [12]. We give a bound on the rate of convergence which is valid for a large region in the parameter space. We use the results to derive a stability and instability theorem for the G.R.C. (Goodwin, Ramadge and Caines [12]) scheme in the presence of unmodeled dynamics and output disturbances and interpret them in terms of frequency contents. The application of these results is substantially different from our earlier work [6,9,10]. Other authors [1,18] have also studied discrete time averaging theory, however, their study has been limited to the linearized systems. Our techniques are general enough to handle more complicated nonlinear discrete systems. As far as the application to adaptive systems is concerned Praly et al and Riedle & Kokotovic consider either the case with periodic signals [24,26] or the model reference adaptive control scheme(Narendra-Lin scheme) [28].

The detailed outline of our paper is as follows: In section 2, we develop a new discrete time converse stability theorem which is a counterpart of a well known continuous time theorem [14,32] and use this to establish the discrete time averaging theorems for systems with one or two time scales. Section 3 applies these theorems to adaptive identification and control schemes. In section 4, we apply averaging to the study of the robustness of discrete adaptive control schemes (largely sampled systems). We first discuss the effect of sampling on the unmodeled dynamics of a continuous time system in section 4.1. In section 4.2, we discuss how when the adaptive identifier has insufficient or bad data for identification that the overall system can slowly drift to instability.

*Research supported by NASA under grant NAG 2-243 and Army Research Office under grant DAAG-29-85-k-0072

2 Converse Lyapunov Theorems and Averaging Theory

2.1 Lyapunov Stability and Input-Output Stability

In recent years, several stability theorems have been developed for the analysis of continuous time systems. However, results for discrete time systems are not as readily available as their continuous counterparts. It is our belief that results which have been developed for continuous systems could equally well be carried through for discrete systems. Of course, some parallels are fairly evident, while others are not.

In this section, we present several results concerning the relationship between input-output and exponential stability of discrete time systems. These results are well-known for continuous time systems [14,32].

Consider a discrete time system described by

$$x(k+1)=f(k,x(k),u(k)) \tag{2.1.1}$$

and

$$x(k+1)=f(k,x(k),0) \tag{2.1.2}$$

where $x\in R^n$,$u\in R^m$ denote the output and input of the system respectively. In what follows, we assume that f(k,0,0)=0 and f has continuous and bounded first order derivative in x. Further, we assume that f satisfies the Lipschitz condition

$$||f(k,x_1,u_1)-f(k,x_2,u_2)|| \leq l_x||x_1-x_2||+l_u||u_1-u_2||$$

$\forall \ k\in Z_+$.

Theorem 2.1.1 (The Converse Lyapunov Theorem)

Consider the system (2.1.2), then the following two statements are equivalent.

(1) x=0 is an exponentially stable equilibrium point of the system (2.1.2), i.e. there exists M> 0 and 0< r< 1 such that

$$||x(k+n)||^2\leq Mr^n||x(k)||^2 \quad \forall \ k,n\in Z_+ \tag{2.1.3}$$

(2) There exists a Lyapunov function $v(x,k)$, and some positive constants α_1 ,α_2 ,α_3 and α_4 such that

$$\alpha_1||x(k)||^2\leq \ v(x(k),k)\leq \ \alpha_2||x(k)||^2 \tag{2.1.4}$$

$$\Delta v_k= \ v(x(k+1),k+1)-v(x(k),k)\leq \ -\alpha_3||x(k)||^2 \tag{2.1.5}$$

$$||\frac{\partial v(x(k),k)}{\partial x(k)}||\leq \ \alpha_4||x(k)|| \tag{2.1.6}$$

Proof:

2) to 1)

This direction is straightforward.

1) to 2)

Let $S_{k+l}(x(k),k)$ denote the solution of the system (2.1.2) at time $(k+l)$ with initial time k and initial value x(k). Define

$$v(x(k),k)=\sum_{i=0}^{n-1}||S_{k+i}(x(k),k)||^2 \tag{2.1.7}$$

where $n\in Z_+$ and satisfies $1-Mr^n>0$. It is obvious that $S_k(x(k),k)=x(k)$ and consequently $v(x(k),k)\geq ||x(k)||^2$.

Note that

$$\begin{aligned} v(x(k),k)&\leq Mr^{n-1}||x(k)||^2+...+Mr||x(k)||^2+||x(k)||^2 \\ &= Mr(r^{n-2}+...+1)||x(k)||^2+||x(k)||^2 \\ &= (Mr\frac{1-r^{n-1}}{1-r}+1)||x(k)||^2 \end{aligned} \tag{2.1.8}$$

and

$$\begin{aligned} \Delta v_k &= ||S_{k+n}(x(k+1),k+1)||^2-||x(k)||^2 \\ &\leq Mr^n||x(k)||^2-||x(k)||^2=-(1-Mr^n)||x(k)||^2 \end{aligned} \tag{2.1.9}$$

This gives (2.1.4) and (2.1.5) with $\alpha_1=1, \alpha_2=Mr\frac{1-r^{n-1}}{1-r}+1$ and $\alpha_3=(1-Mr^n)$. The proof of the last inequality is by direct calculation. Observe that

$$v(x(k),k)=x^T(k)x(k)+\sum_{i=0}^{n-2}f^T(k+i,x(k+i),0)f(k+i,x(k+i),0)$$

$$\frac{\partial v(x(k),k)}{\partial x(k)}=2x(k)+2(\frac{\partial f(k,x(k),0)}{\partial x(k)})^Tx(k+1)...$$

$$+2((\frac{\partial f(k+n-2,x(k+n-2),0)}{\partial x(k+n-2)})...(\frac{\partial f(k,x(k),0)}{\partial x(k)}))^Tx(k+n-1)$$

Since f has continuous and bounded first order partial derivative, the conclusion follows.

$$||\frac{\partial v}{\partial x(k)}||\leq \alpha_4||x(k)||$$

for some $\alpha_4>0$.

Q.E.D.

Remark:

Theorem 2.1 is a global property. However, if the assumptions are valid locally and we assume also that x=0 is a stable equilibrium point, then the conclusion holds locally.

Theorem 2.1.2 (Input-Output Finite Gain l_∞ Stability)

Consider the systems (2.1.1) and (2.1.2). Assume that x=0 is an exponentially stable equilibrium point of the system (2.1.2), then the system (2.1.1) is input-output finite gain l_∞ stable, i.e. there exists a finite constant r such that whenever $u \in l_\infty^m$ i.e. $\|u(\cdot)\|_\infty = U_m < \infty$,the resulting $x \in l_\infty^n$ and satisfies

$$\|x(\cdot)\|_\infty \le r\,\|u(\cdot)\|_\infty$$

Proof: The hypothesis implies that there exists a Lyapunov function v for the system (2.1.2) such that inequalities (2.1.4), (2.1.5) and (2.1.6) are satisfied. If we evaluate the difference of v along the trajectory of system (2.1.1), we get

$$\begin{aligned} v(x(k+1),k+1)-v(x(k),k) = {} & v(f(k,x(k),u(k)),k+1)-v(f(k,x(k),0),k+1) \\ & + v(f(k,x(k),0),k+1)-v(x(k),k) \\ & \le \alpha_4\|\lambda f(k,x(k),u(k))+(1-\lambda)f(k,x(k),0)\| \\ & \cdot \|f(k,x(k),u(k))-f(k,x(k),0)\|-\alpha_3\|x(k)\|^2 \\ & \le \alpha_4(l_u\|u(k)\|+l_x\|x(k)\|)l_u\|u(k)\|-\alpha_3\|x(k)\|^2 \end{aligned} \tag{2.1.10}$$

for some $\lambda \in [0,1]$, where in the second step we use inequalities (2.1.5) and (2.1.6) and also the mean value theorem. This implies that

$$\Delta v_k \le M_1 + M_2\|x(k)\| - \alpha_3\|x(k)\|^2$$

with $M_1 = \alpha_4 l_u^2 U_m^2, M_2 = \alpha_4 l_x l_u U_m$. From this, we can say that $\|x(k)\|$ is bounded for any k. In fact,

$$\Delta v_k \le M_1 + M_2\left(\frac{v(x(k),k)}{\alpha_1}\right)^{1/2} - \frac{\alpha_3 v(x(k),k)}{\alpha_2} \tag{2.1.11}$$

Observe now that the right hand side of (2.1.11) is negative whenever $v(x(k),k) > v_0$, where v_0 is a solution of the equation

$$M_1 + M_2\left(\frac{v}{\alpha_1}\right)^{1/2} - \frac{\alpha_3 v}{\alpha_2} = 0$$

i.e.

$$v_0^{1/2} = \frac{\alpha_2}{2\alpha_3}\left(\frac{\alpha_4 l_x l_u}{\alpha_1^{1/2}} + \left(\frac{\alpha_4^2 l_x^2 l_u^2}{\alpha_1} + \frac{4\alpha_3\alpha_4 l_u^2}{\alpha_2}\right)^{1/2}\right)U_m = r_1 U_m \tag{2.1.12}$$

Now from (2.1.4), we have that

$$v(x(k+1),k+1)\leq \alpha_2\|x(k+1)\|^2\leq \alpha_2(l_x\|x(k)\|+l_u\|u(k)\|)^2$$

$$\leq \alpha_2\left(l_x\left(\frac{v(x(k),k)}{\alpha_1}\right)^{1/2}+l_u\|u(k)\|\right)^2$$

Using the maximum value v_0 for $v((x(k),k)$, we have

$$\text{Max } v\leq \alpha_2\left(l_x\left(\frac{v_0}{\alpha_1}\right)^{1/2}+l_u U_m\right)^2 = r_2 U_m{}^2 \tag{2.1.13}$$

The conclusion then follows from (2.1.13).

$$\|x(k)\|\leq \left(\frac{\text{Max } v}{\alpha_1}\right)^{1/2} = rU_m \qquad \forall \;\; k\in Z_+ \tag{2.1.14}$$

Q.E.D.

Corollary 2.3

Consider the system (2.1.1) and (2.1.2) with all assumptions of theorem 2.2 being satisfied. Moreover assume that

$$\|u(k)\|\leq \delta\|x(k)\| \qquad \forall \;\; k\in Z_+$$

with $\alpha_3-\alpha_4 l_u{}^2\delta^2-\alpha_4 l_u l_x\delta>0$. Then the system (2.1.1) is also exponentially stable.

Proof: Follows immediately from (2.1.10).

Q.E.D.

2.2 Averaging Theory

In this section, we will give several discrete time averaging theorems using a converse Lyapunov stability theorem. The proofs are omitted (however available from authors upon request) since they are similar to those in the continuous time cases [6,9,10].

Consider the difference equation of the form

$$x(k+1)=x(k)+\epsilon f(k,x(k),\epsilon) \tag{2.2.1}$$

where $x\in R^n$, $k\in Z_+$, $0<\epsilon\leq\epsilon_0$ and f is piecewise continuous in k with the limit

$$f_{av}(x)=\lim_{T\to\infty}\frac{1}{T}\sum_{k=s+1}^{s+T} f(k,x,0) \tag{2.2.2}$$

existing uniformly in s and $\forall\; x\in B_h$, a closed ball in R^n of radius h. Assume that f and f_{av} satisfy the following conditions ($\forall\; x\in B_h$, $0<\epsilon\leq\epsilon_0$ and $k\in Z_+$).

(A1) x=0 is an equilibrium point of the equation (2.2.1) i.e. $f(k,0,\epsilon)=0$ and f is Lipschitz in x

$$\|f(k,x_1,\epsilon)-f(k,x_2,\epsilon)\| \le l_1\|x_1-x_2\| \tag{2.2.3}$$

(A2) $f(k,x,\epsilon)$ is Lipschitz in ϵ, linearly in x

$$\|f(k,x,\epsilon_1)-f(k,x,\epsilon_2)\| \le l_2|\epsilon_1-\epsilon_2|\,\|x\| \tag{2.2.4}$$

(A3) $f_{av}(0)=0$ and $f_{av}(x)$ is Lipschitz in x

$$\|f_{av}(x_1)-f_{av}(x_2)\| \le l_{av}\|x_1-x_2\| \tag{2.2.5}$$

(A4) Define $d(k,x)=f(k,x,0)-f_{av}(x)$. Then d(k,x) is piecewise continuous in k and has bounded and continuous first partial derivative in x and d(k,0)=0. Moreover, there is a non-negative strictly decreasing function $\gamma(k)$ with the property $\gamma(k)\to 0$ as $k\to\infty$, so that

$$\|\frac{1}{T}\sum_{k=s+1}^{s+T} d(k,x)\| \le \gamma(T)\|x\| \tag{2.2.6}$$

and

$$\|\frac{1}{T}\sum_{k=s+1}^{s+T} \frac{\partial d(k,x)}{\partial x}\| \le \gamma(T) \tag{2.2.7}$$

The averaged system of the equation (2.2.1) is defined as

$$x_{av}(k+1)=x_{av}(k)+\epsilon f_{av}(x_{av}(k)) \tag{2.2.8}$$

Then we have following theorems.

Theorem 2.2.1 (Basic Averaging Theorem)

Consider the original system (2.2.1) and the averaged system (2.2.8) satisfying the assumptions (A1)-(A4). For any given $T\in Z_+$, further assume that the initial condition x_0 is sufficient small so that $x_{av}(k)\in B_{h'}$ for some $h' < h$ and $k\in[0,[T/\epsilon]]$ (where $[T/\epsilon]$ denotes the largest integer l such that $l \le T/\epsilon$). Then there is an $\epsilon_T, 0<\epsilon_T\le\epsilon_0$ and a class K function $\psi(\epsilon)$ (i.e. a positive nondecreasing function and $\psi(0)=0$), so that

$$\|x(k)-x_{av}(k)\| \le \psi(\epsilon)b_T \tag{2.2.9}$$

for some $b_T>0$, all $k\in[0,[T/\epsilon]]$ and $0<\epsilon\le\epsilon_T$.

Theorem 2.2.2 (Exponential Stability Theorem)

Consider the original system (2.2.1) and the averaged system (2.2.8) satisfying the assumptions (A1)-(A4). Further assume that $f_{av}(x)$ has continuous and bounded first partial derivative in x. If x=0 is an exponentially stable equilibrium point of the averaged system (2.2.8), then there exists an ϵ_2, $0<\epsilon_2\le\epsilon_0$, such that x=0 is an exponentially stable equilibrium point of the original system (2.2.1) for all $0<\epsilon\le\epsilon_2$.

Now, let us consider two-time scale systems of the form

$$x(k+1) = x(k) + \epsilon f(k,x(k),y(k),\epsilon) \tag{2.2.10}$$

$$y(k+1) = A(x(k))y(k) + \epsilon g(k,x(k),y(k),\epsilon) \tag{2.2.11}$$

where $x \in R^n$ is the slow state and $y \in R^m$ is the fast state.

The averaged system of (2.2.10) is defined by

$$x_{av}(k+1) = x_{av}(k) + \epsilon f_{av}(x_{av}(k)) \tag{2.2.12}$$

where f_{av} is the limit (assuming this limit exists uniformly in s and $\forall\ x \in B_h$)

$$f_{av}(x) = \lim_{T \to \infty} \frac{1}{T} \sum_{k=s+1}^{s+T} f(k,x,0,0) \tag{2.2.13}$$

Assuming that f and g satisfy the following assumptions ($\forall\ x \in B_h,\ y \in B_h,\ 0 < \epsilon \leq \epsilon_0$ and $k \in Z_+$).

(B1) x=0,y=0 is an equilibrium point of the system (2.2.10) and (2.2.11), i.e.

$$f(k,0,0,\epsilon) = 0, \quad g(k,0,0,\epsilon) = 0$$

and f and g are Lipschitz in x and y

$$\|f(k,x_1,y_1,\epsilon) - f(k,x_2,y_2,\epsilon)\| \leq l_1\|x_1 - x_2\| + l_2\|y_1 - y_2\|$$

$$\|g(k,x_1,y_1,\epsilon) - g(k,x_2,y_2,\epsilon)\| \leq l_3\|x_1 - x_2\| + l_4\|y_1 - y_2\|$$

(B2) $f(k,x,y,\epsilon)$ and $g(k,x,y,\epsilon)$ are Lipschitz in ϵ, linearly in x,y

$$\|f(k,x,y,\epsilon_1) - f(k,x,y,\epsilon_2)\| \leq l_5(\|x\| + \|y\|)\,|\epsilon_1 - \epsilon_2|$$

$$\|g(k,x,y,\epsilon_1) - g(k,x,y,\epsilon_2)\| \leq l_6(\|x\| + \|y\|)\,|\epsilon_1 - \epsilon_2|$$

(B3) $f_{av}(0) = 0$ and $f_{av}(x)$ is Lipschitz in x

$$\|f_{av}(x_1) - f_{av}(x_2)\| \leq l_{av}\|x_1 - x_2\|$$

(B4) Let $d(k,x) = f(k,x,0,0) - f_{av}(x)$. Then d(k,x) should satisfy the assumption (A4).

(B5) $A(x) \in R^{m \times m}$ is uniformly exponentially stable $\forall\ x \in B_h$, i.e. there exist m_1 , $m_2 > 0$ and λ_1 , $\lambda_2 \in [0,1)$ such that

$$m_1 \lambda_1^k \leq \|A(x)^k\| \leq m_2 \lambda_2^k \qquad \forall\ x \in B_h \tag{2.2.14}$$

Moreover

$$\left\|\frac{\partial A(x)}{\partial x_i}\right\| \leq k_a \qquad i = 1,..,n \tag{2.2.15}$$

for some $K_a > 0$

(B6) x_0 and y_0 are small enough that $x_{av} \in B_{h'}$ for some $h' < h$ on the time interval considered (to be specified shortly).

Remark:

Assumption (B5) implies that there exists a symmetric matrix $P(x) \in R^{m \times m}$ and some constants $p_1, p_2, q_1, q_2 > 0$ such that

$$p_1 I \leq P(x) \leq p_2 I$$

$$-q_1 I \leq A^T(x)P(x) + P(x)A(x) \leq -q_2 I$$

Theorem 2.2.3 (Basic Averaging Theorem for Two-Time Scale Systems)

Consider the system (2.2.10) and (2.2.11) and the averaged system (2.2.12) satisfying the assumptions (B1)-(B6). Then for any given $T \in Z_+$, there exists $b_T > 0$, $0 < \epsilon_T \leq \epsilon_0$ and a class K function $\psi(\epsilon)$ such that

$$||x(k) - x_{av}(k)|| \leq \psi(\epsilon) b_T \qquad (2.2.16)$$

Theorem 2.2.4: (Exponential Stability Theorem for Two-Time Scale System)

Consider the original system (2.2.10) and (2.2.11) and the averaged system (2.2.12) satisfying the assumptions (B1)-(B6). If the averaged system (2.2.12) is exponentially stable, then there exists $0 < \epsilon_1 \leq \epsilon_0$ such that the original system (2.2.10) and (2.2.11) is exponentially stable for all $0 < \epsilon \leq \epsilon_1$.

To end this section, let us discuss a mixed time scale system of the form

$$x(k+1) = x(k) + \epsilon f'(k, x(k), y'(k), \epsilon) \qquad (2.2.17)$$

$$y'(k+1) = A(x(k))y'(k) + h(k, x(k)) + \epsilon g'(k, x(k), y'(k), \epsilon) \qquad (2.2.18)$$

where h satisfies

(B7) h (k,0)=0 and

$$||\frac{\partial h(k,x)}{\partial x}|| \leq k_h$$

for some $k_h > 0$ and $\forall\ x \in B_h$.

By defining the function,

$$w(k,x) = \sum_{i=0}^{k-1} A(x)^{k-i-1} h(i,x) \qquad (2.2.19)$$

we may construct the transformation

$$y(k) = y'(k) - w(k,x) \qquad (2.2.20)$$

to transform the mixed time scale system (2.2.17) and (2.2.18) into the form of the two time scale system

$$x(k+1) = x(k) + \epsilon f(k, x(k), y(k), \epsilon) \tag{2.2.21}$$

$$y(k+1) = A(x(k))y(k) + \epsilon g(k, x(k), y(k), \epsilon) \tag{2.2.22}$$

with

$$f(k,x,y,\epsilon) = f'(k,x,y+w(k,x),\epsilon) \tag{2.2.23}$$

and

$$g(k,x,y,\epsilon) = g'(k,x,y+w(k,x),\epsilon)$$
$$-\left\{ \int_0^1 \frac{\partial w}{\partial x}(k+1, sx(k+1)+(1-s)x(k))\,ds \right\} f'(k,x,y+w(k,x),\epsilon) \tag{2.2.24}$$

Assuming that assumptions (B1)-(B6) are satisfied, the two-time scale averaging theorems may be applied to mixed-time scale systems.

3 Applications of Averaging Theory to Adaptive Systems

3.1 The Adaptive Identifier

In this section, we apply averaging theory to adaptive identifiers, particularly to the convergence analysis of the projection type algorithm and the least squares type algorithm.

Consider a discrete time system in the standard form [12]

$$y(k)=\phi^T(k-1)\theta_0 \tag{3.1.1}$$

where $\theta_0{}^T=(a_1,..,a_n,b_1,..,b_m)$ is an unknown parameter vector with n,m known and $\phi^T(k-1)=(y(k-1),..,y(k-n),u(k-1),..u(k-m))$ the regressor vector.

For the identifier using the projection type algorithm

$$\theta(k)=\theta(k-1)+\frac{\epsilon\phi(k-1)}{1+\epsilon\phi^T(k-1)\phi(k-1)}(y(k)-\phi^T(k-1)\theta(k-1)) \tag{3.1.2}$$

The parameter error equation may be written as

$$\tilde{\theta}(k)=\tilde{\theta}(k-1)-\frac{\epsilon\phi(k-1)\phi^T(k-1)}{1+\epsilon\phi^T(k-1)\phi(k-1)}\tilde{\theta}(k-1) \tag{3.1.3}$$

with $\tilde{\theta}(k)=\theta(k)-\theta_0$ denoting the parameter error. To apply averaging theory developed in the previous section, we need the existence of the averaged system of equation (3.1.3). For this purpose, we assume that the system (3.1.1) is stable and the input u(k) is stationary that the following limit exists uniformly in s

$$\lim_{T\to\infty}\frac{1}{T}\sum_{i=s+1}^{s+T}\phi(i)\phi^T(i)=R \tag{3.1.4}$$

Then, the averaged system of (3.1.3) is defined by (for small $\epsilon>0$)

$$\tilde{\theta}_{av}(k)=(I-\epsilon R)\tilde{\theta}_{av}(k-1) \tag{3.1.5}$$

Denoting by $\bar{\sigma}(A)$ ($\underline{\sigma}(A)$) the maximum (minimum) singular value (or eigenvalue) of matrix A, we observe that

$$\bar{\sigma}(I-\epsilon R)=1-\epsilon\underline{\sigma}(R) \tag{3.1.6}$$

Thus by the averaging results of the last section, we may conclude that the convergence rate of the projection algorithm (3.1.2) is faster than $(1-\epsilon\underline{\sigma}(R))$ (up to high order terms in ϵ) for small $\epsilon>0$. i.e.

$$||\tilde{\theta}(k)||\le M(1-\epsilon\underline{\sigma}(R))^k\,||\tilde{\theta}(0)||$$

for some $M>0$.

Now, we relate the convergence rate to the spectrum of the input. Notice that the input-output relation can be written as

$$y(z) = G(z)u(z) \tag{3.1.7}$$

consequently,

$$\phi(z) = \begin{bmatrix} G(z) \\ z^{-1}G(z) \\ \cdot \\ \cdot \\ z^{-n+1}G(z) \\ 1 \\ z^{-1} \\ \cdot \\ \cdot \\ z^{-m+1} \end{bmatrix} u(z) = n(z)u(z) \tag{3.1.8}$$

From the Herglotz Theorem, we may write R in terms of the input spectrum and the transfer function n(z) as

$$R = \frac{1}{2\pi}\int_{-\pi}^{\pi} n(e^{jw})n^T(e^{-jw})S_u(dw) \tag{3.1.9}$$

where $S_u(dw)$ stands for the spectral measure of the input u. For optimum convergence, the spectrum of the input needs to lie in the region where $\underline{\sigma}(R)$ is large.

For the analysis of the slowed-down least squares type algorithm, we write the algorithm as follows (for small $\epsilon > 0$)

$$\theta(k) = \theta(k-1) - \epsilon P(k-1)\phi(k-1)(y(k) - \phi^T(k-1)\theta(k-1)) \tag{3.1.10a}$$

$$P^{-1}(k-1) = P^{-1}(k-2) - \epsilon(\rho P^{-1}(k-2) + \phi(k-1)\phi^T(k-1)) \tag{3.1.10b}$$

where $0 < \rho < 1$. As before, we may approximate equation (3.1.10) by its averaged system (assume that $u(k)$ is sufficiently rich so that R is nonsingular (see Bai & Sastry [5]))

$$\tilde{\theta}_{av}(k) = (1 - \epsilon P_{av}(k-1)R)\tilde{\theta}_{av}(k-1) \tag{3.1.11a}$$

$$P_{av}^{-1}(k) = P_{av}^{-1}(k-1) + \epsilon(\rho P_{av}^{-1}(k-1) + R) \tag{3.1.11b}$$

R is defined in (3.1.4). Equation (3.1.11b) may be explicitly solved to give

$$P_{av}^{-1}(k) = (1-\epsilon\rho)^k P_{av}^{-1}(0) + R\frac{1}{\rho}(1-(1-\epsilon\rho)^k) \tag{3.1.12}$$

In turn, using this in equation (3.1.11a) and noting that $P_{av}^{-1}(k)$ converges to $R\frac{1}{\rho}$ as $t \to \infty$. we see that the tail behavior of (3.1.11a) is

$$\tilde{\theta}_{av}(k) = (1-\epsilon\rho)\tilde{\theta}_{av}(k-1) \tag{3.1.13}$$

so that the tail convergence rate is a function of ρ and gain ϵ alone and not the input spectrum!

3.2 Application to Adaptive Control

We apply the averaging theory to the discrete-time adaptive control scheme of Goodwin, Ramadge and Caines (Our results are extensions of Praly [22-24,26]).

Consider a plant modeled by

$$A(q^{-1})\,y(k) = q^{-d}B(q^{-1})\,u(k) \tag{3.2.1}$$

where

$$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n} \tag{3.2.2}$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + \cdots + b_m q^{-m} \tag{3.2.3}$$

and d is a pure time delay. To facilitate direct control, the model (3.2.1) is converted into its d-step-ahead predictor form

$$y(k+d) = \alpha(q^{-1})\,y(k) + \beta(q^{-1})\,u(k) \tag{3.2.4}$$

where

$$\alpha(q^{-1}) = \alpha_0 + \alpha_1 q^{-1} + \cdots \alpha_{n-1} q^{-(n-1)} \tag{3.2.5}$$

$$\beta(q^{-1}) = \beta_0 + \beta_1 q^{-1} + \cdots \beta_{m+d-1} q^{-(m+d-1)} \tag{3.2.6}$$

and

$$\beta_0 = b_0 \neq 0 \tag{3.2.7}$$

The objective of the adaptive control scheme is to get $y(k)$ to track a given reference trajectory $y^*(k)$. The conditions on $A(q^{-1})$ and $B(q^{-1})$ are identical to those in [12].

Denoting

$$\theta_0^T = (\alpha_0, \alpha_1, \cdots, \alpha_{n-1}, \beta_0, \cdots, \beta_{m+d-1}) \tag{3.2.8}$$

(3.2.4) can be written as

$$y(k+d) = \phi^T(k)\cdot\theta_0 \tag{3.2.9}$$

where

$$\phi(k)^T = (\,y(k), \cdots, y(k-n+1), u(k), \cdots, u(k-m-d+1)\,) \tag{3.2.10}$$

The control law is implemented as:

$$y^*(k+d) = \phi^T(k)\cdot\theta(k) \tag{3.2.11}$$

where $\theta(k)$ denotes an estimate of θ_0 at time k.

Goodwin, Ramadge and Caines have shown that, under the projection type parameter adaptation law:

$$\theta(k) = \theta(k-1) + \epsilon \frac{\phi(k-d)(y(k) - \phi^T(k-d)\theta(k-1))}{1 + \epsilon\phi^T(k-d)\phi(k-d)} \tag{3.2.12}$$

the following are true.

(i) {y(k)} and {u(k)} are bounded sequences.

(ii) $\lim_{k \to \infty} [y(k) - y^*(k)] = 0.$

(iii) if $\phi(k)$ is persistently exciting, then $(\theta(k) - \theta_0) \to 0$ exponentially.

To facilitate the averaging analysis of such adaptive control system, we convert the system (3.2.1), (3.2.11) and (3.2.12) into its state space form (as in [24,26]).

$$x(k+1) = A(\theta(k))x(k) + B(\theta(k))y^*(k+d) \tag{3.2.13}$$

$$y(k) = c^T x(k) \tag{3.2.14}$$

and

$$\theta(k) = \theta(k-1) + \epsilon f'(x(k-1), \theta(k-1), \epsilon) \tag{3.2.15}$$

where $x(k)$ is the state vector containing all the elements of $\phi(k-d+1)$ and satisfying $\phi(k-d+1) = Q\,x(k)$ for a suitable constant matrix Q.

As implied by the fact (ii), there exists a compact set $B(\theta_0, h)$ in R^{m+n+d} such that $A(\theta_0)$ is uniformly exponentially stable as in (2.2.14). Define the state error $\tilde{x}(k)$ and parameter error $\tilde{\theta}(k)$ respectively by

$$\tilde{x}(k) = x(k) - x_{\theta_0}(k) \tag{3.2.16}$$

$$\tilde{\theta}(k) = \theta(k) - \theta_0 \tag{3.2.17}$$

where $x_{\theta_0}(k)$ is the state that would be obtained if $\theta(k)$ were frozen at θ_0. Obviously, $y^*(k) = c^T x_{\theta_0}(k)$. Now the error model can be formulated as:

$$\tilde{x}(k+1) = A(\theta)\tilde{x}(k) + h(k, \tilde{\theta}(k)) \tag{3.2.18}$$

$$\tilde{\theta}(k) = \tilde{\theta}(k-1) + \epsilon f(k, \tilde{x}(k-1), \tilde{\theta}(k-1), \epsilon) \tag{3.2.19}$$

where

$$h(k, \tilde{\theta}(k)) = [A(\tilde{\theta}(k) + \theta_0) - A(\theta_0)]\, x_{\theta_0}(k)$$
$$+ [B(\tilde{\theta}(k) + \theta_0) - B(\theta_0)]\, y^*(k+d) \tag{3.2.20}$$

$$f(k, \tilde{x}(k-1), \tilde{\theta}(k-1), \epsilon) = f'(\tilde{x}(k-1) + x_{\theta_0}(k-1), \tilde{\theta}(k-1) + \theta_0, \epsilon) \tag{3.2.21}$$

satisfying assumptions (B1)-(B7) in section 2 for all $\epsilon \leq \epsilon_2$, $\epsilon_2 > 0$. System (3.2.18), (3.2.19) appear in the same form as those in (2.2.17) and (2.2.18) which allows for the direct application of mixed-time scale averaging results, namely, the slow state $\tilde{\theta}(k)$ can be approximated by its averaged system

$$\tilde{\theta}_{av}(k+1) = \tilde{\theta}_{av}(k) + \epsilon f_{av}(\tilde{\theta}_{av}(k)) \tag{3.2.22}$$

where f_{av} is defined by the limit

$$f_{av}(\theta) = \lim_{T\to\infty} \frac{1}{T} \sum_{i=s+1}^{s+T} f(i, w(i,\theta), \theta, 0) \tag{3.2.23}$$

and

$$w(k,\theta) = \sum_{i=0}^{k-1} A(\theta)^{k-i-1} h(i,\theta) \tag{3.2.24}$$

To simplify the expression of $f_{av}(\theta)$ in (3.2.23), we assume that the reference trajectory signal $y^*(k)$ is stationary. Referring to (3.2.12), f can be expressed as:

$$f(k, \tilde{x}(k-1), \tilde{\theta}(k-1), \epsilon) = -\frac{\phi(k-d)\phi^T(k-d)\tilde{\theta}(k-1)}{1+\epsilon\phi^T(k-d)\phi(k-d)} \tag{3.2.25}$$

and (3.2.24) implies $\phi(k-d)$ may be replaced by $\phi_\theta(k-d)$, a regressor signal that would be obtained if parameter $\theta(k)$ were fixed at θ, which results in the following expression of $f_{av}(\theta)$:

$$f_{av}(\tilde{\theta}) = -\left\{ \lim_{T\to\infty} \frac{1}{T} \sum_{k=s+1}^{s+T} \phi_\theta(k-d)\phi^T_\theta(k-d) \right\}\tilde{\theta} \qquad s \geq d-1 \tag{3.2.26}$$

Using a little algebra, one can show that

$$\phi_\theta(k) = \frac{C(q^{-1})}{C(q^{-1})^T \cdot \theta} y^*(k+d) \tag{3.2.27}$$

where

$$C^T(q^{-1}) = (q^{-d}B(q^{-1}), \cdots, q^{-(n+d-1)}B(q^{-1}), A(q^{-1}), \cdots, q^{-(m+d-1)}A(q^{-1}))$$

Consequently, by Herglotz's theorem, $f_{av}(\tilde{\theta})$ can be reexpressed as:

$$f_{av}(\tilde{\theta}) = -\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{C(e^{-j\omega})\, C^T(e^{j\omega})}{|\, C^T(e^{-j\omega})\cdot\theta\,|^2} S_{y^*}(d\omega) \right\}\tilde{\theta}$$

$$:= -R_\theta(0)\,\tilde{\theta} \qquad (\theta = \tilde{\theta} + \theta_0) \tag{3.2.28}$$

If $y^*(k)$ contains only finite number of spectral lines, then the integral above is replaced by summation, i.e.

$$R_\theta(0) = \frac{1}{2\pi} \sum_{i=1}^{n_0} \frac{C(e^{-j\omega_i})\, C^T(e^{j\omega_i})}{|\, C^T(e^{-j\omega_i})\cdot\theta\,|^2} r_i^2 \tag{3.2.29}$$

Similarly, averaging results can also be applied to the adaptive scheme with a least squares type adaptation algorithm which is of the form

$$\tilde{\theta}(k) = \tilde{\theta} + \epsilon \frac{P(k-d-1)\phi(k-d)(y(k)-\theta^T(k-1)\cdot\phi(k-d))}{1+\epsilon\phi^T(k-d)P(k-d-1)\phi(k-d)} \tag{3.2.30}$$

$$P(k-d) = \frac{1}{\rho}\left\{P(k-d-1)-\epsilon\frac{P(k-d-1)\phi(k-d)\phi^T(k-d)P(k-d-1)}{\rho+\epsilon\phi^T(k-d)P(k-d-1)\phi(k-d)}\right\} \tag{3.2.31}$$

where $0<\rho<1$. Direct application of averaging to the system (3.2.30) and (3.2.31) is almost the same as before with the only difference being that $\tilde{\theta}(k)$ and $P(k)$ now constitute the slow state variable. As a result, the averaged system of (3.2.30), (3.2.31) can be formulated as:

$$\tilde{\theta}_{av}(k) = \tilde{\theta}_{av}(k-1)-\epsilon P_{av}(k-d-1)R_{\theta_{av}}(0)\tilde{\theta}_{av}(k-1) \tag{3.2.32}$$

$$P_{av}(k-d) = \frac{1}{\rho}\left\{P_{av}(k-d-1)-\frac{\epsilon}{\rho}P_{av}(k-d-1)R_{\theta_{av}}(0)P_{av}(k-d-1)\right\} \tag{3.2.33}$$

when $R_\theta(0)$ is as defined in (3.2.29).

To study the tail behavior of $\tilde{\theta}_{av}(k)$ and $P_{av}(k)$, one can see that $P_{av}(k-d)$ will be close to $\rho R_{\theta_0}(0)^{-1}$ when $\tilde{\theta}_{av}$ is sufficiently close to zero, while the slowly evolving dynamics of $\tilde{\theta}_{av}$ becomes:

$$\begin{aligned}\tilde{\theta}_{av}(k) &\approx \tilde{\theta}_{av}(k-1)+\epsilon(\rho R_{\theta_0}(0)^{-1})\cdot(-R_{\theta_0}(0))\tilde{\theta}_{av}(k-1)\\ &= (1-\epsilon\rho)\tilde{\theta}_{av}(k-1)\end{aligned} \tag{3.2.34}$$

Remark:

It has been shown in [5] that the parameter error will converge to zero exponentially if $y^*(k)$ contains no less than $(n+m+d)$ spectral lines. For the projection type adaptation algorithm, the convergence rate of the averaged system (3.2.22) is guaranteed to be at least $(1-\epsilon\alpha_2)$ where

$$\alpha_2 \le \lambda[R_\theta(0)] \qquad \forall\ \theta\in B(\theta_0,h) \tag{3.2.35}$$

The rate of convergence of the original system can be estimated by the same value for ϵ sufficiently small. It is interesting to note that, from (3.2.29), as $\|\theta\|$ increases, $\lambda_{\min}(R_\theta(0))$ tends to zero in some directions. This indicates that the adaptive control system is not globally exponentially stable (with uniform convergence rate).

For the least squares type adaptation, the tail convergence rate of $\tilde{\theta}_{av}(k)$ doesn't rely on the exogenous reference trajectory (when $\|\tilde{\theta}_{av}\|$ is small enough) as indicated in (3.2.34).

4 Averaging Analysis of Robustness in The Presence of Unmodeled Dynamics and Bounded Output Disturbances

4.1 Unmodeled Dynamics in Discrete and Sampled Systems

Unmodeled dynamics in continuous time systems have been well documented (see, for instance, Doyle & Stein [8] for a good discussion). They are usually assumed to be of the form

$$\overline{G}(s)= G(s)(1+\Delta G_1(s))+\Delta G_2(s) \tag{4.1.1}$$

where $\overline{G}(s)$ and $G(s)$ are true and nominal systems. $\Delta G_1(s)$ and $\Delta G_2(s)$ are called additive uncertainty and multiplicative uncertainty respectively. For most adaptive systems, it has been further assumed that $\Delta G_1(s)$ and $\Delta G_2(s)$ are finite order, proper stable rational functions. This implies implicitly that the system $\overline{G}(s)$ is a finite order linear time invariant system and the nominal system $G(s)$ captures all unstable modes. Motivated by this, some authors [15,21] define uncertainty in discrete time systems by

$$\overline{H}(z)= H(z)(1+\Delta H_1(z))+\Delta H_2(z) \tag{4.1.2}$$

$H(z)$, $\Delta H_1(z)$ and $\Delta H_2(z)$ denote the nominal system, additive uncertainty and multiplicative uncertainty respectively. As in the continuous time case, $\Delta H_1(z)$ and $\Delta H_2(z)$ are assumed to be finite order, proper stable rational functions in z.

Under these assumptions, the overall system may be drawn as in Fig. 1.

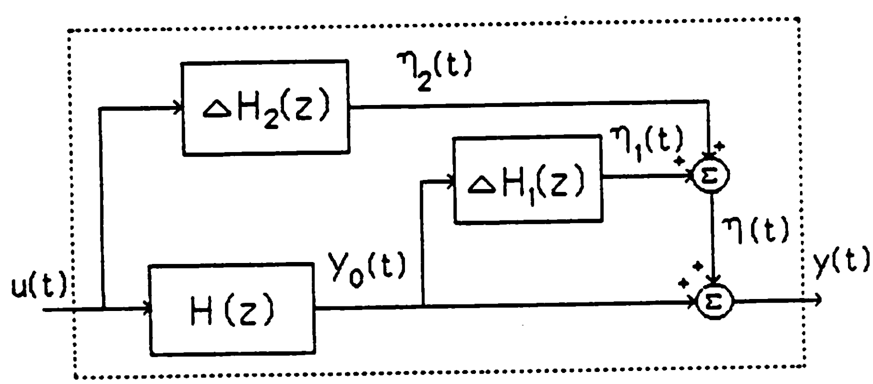

Fig. 1 $\overline{H}(z)$

Define the nominal output $y_0(k)$ by

$$y_0(k) = H(z)u(k) \tag{4.1.3}$$

and the error term $\eta(k)$ by

$$\eta(k) = y(k) - y_0(k) \tag{4.1.4}$$

It follows that

Fact 1.

Consider the discrete system (4.1.2). Suppose that $\Delta H_1(z)$ and $\Delta H_2(z)$ are finite order, proper stable, rational functions in z. Then, the error term $\eta(k)$ satisfies

$$|\eta(k)| \le M(1 + \sum_{i=0}^{k-1} \lambda^{k-1-i}(|u(i)| + |y_0(i)|) + |u(k)| + |y_0(k)|) \tag{4.1.5}$$

for some $M > 0$, $0 < \lambda < 1$ and all $k \in Z_+$.

Proof: Notice that

$$\eta(k) = \eta_1(k) + \eta_2(k) = \Delta H_1(z) y_0(k) + \Delta H_2(z) u(k)$$

Let (A_i, B_i, C_i, D_i) denote the state space realization of $\Delta H_i(z)$ (i=1,2), then $\eta_1(k)$ and $\eta_2(k)$ may be considered as the outputs of the following stable systems

$$x_1(k+1) = A_1 x_1(k) + B_1 y_0(k)$$
$$\eta_1(k) = C_1 x_1(k) + D_1 y_0(k) \tag{4.1.6}$$

and

$$x_2(k+1) = A_2 x_2(k) + B_2 u(k)$$
$$\eta_2(k) = C_2 x_2(k) + D_2 u(k). \tag{4.1.7}$$

Hence

$$|\eta_1(k)| \le |C_1 A_1{}^k x_1(0)| + |C_1 \sum_{i=0}^{k-1} A_1{}^{k-1-i} B_1 y_0(i)| + |D_1 y_0(k)|$$

$$|\eta_2(k)| \le |C_2 A_2{}^k x_2(0)| + |C_2 \sum_{i=0}^{k-1} A_2{}^{k-1-i} B_2 u(i)| + |D_2 u(k)|$$

and this implies that

$$|\eta(k)| \le |\eta_1(k)| + |\eta_2(k)|$$

$$\le M(1 + \sum_{i=0}^{k-1} \lambda^{k-1-i}(|u(i)| + |y_0(i)|) + |u(k)| + |y_0(k)|) \tag{4.1.8}$$

for some $M > 0$ and some $0 < \lambda < 1$. This completes the proof.

Q.E.D.

Remarks:

(1) The bound on $\eta(k)$ is similar to that on the normalizing signal of Praly [25] or the modeling error of Kreisselmeier & Anderson [20]. The difference is that the upper bound on $|\eta(k)|$ here depends not only on $|u(i)|$, $|y_0(i)|$ for $i<k$, but also on the current values of $|u(k)|$ and $|y_0(k)|$. By doing so, the effect of an unmodeled dynamics due to additive and multiplicative uncertainty may be interpreted as a disturbance acting on the nominal output, which depends on the input and the nominal model. Notice that $\eta(k)$ may grow without bound.

(2) If $\Delta H_1(z)$ and $\Delta H_2(z)$ are strictly proper, then (4.1.5) may be modified as

$$|\eta(k)| \le M(1+\sum_{i=1}^{k-1}\lambda^{k-1-i}(|u(i)|+|y_0(i)|)$$

(3) The discrete systems of the form (4.1.2) are not contrived. An example is a near pole-zero cancellation

$$\frac{1}{z+\alpha}\,\frac{z+\beta+\epsilon}{z+\beta}=\frac{1}{z+\alpha}\,(1+\frac{\epsilon}{z+\beta})$$

A large class of discrete systems arises from sampling of continuous systems as shown in Fig.2 (h is the sampling interval.)

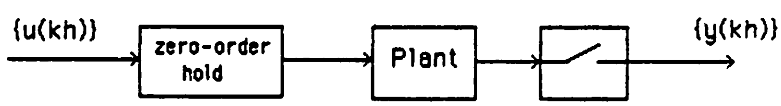

Fig. 2

Consequently, it is important to study the effect of sampling on the uncertainty in continuous time systems.

For the additive uncertainty in continuous systems

$$\overline{G}(s)= G(s)+\Delta G_2(s)$$

the sampled systems become

$$\overline{H}(z)= H(z)+ \Delta H_2(z)$$

with $H(z)$ nominal model (sampling of G(s)) and $\Delta H_2(z)$ additive uncertainty (sampling of $\Delta G_2(s)$). Since the poles p_i of a continuous time system are transformed, under sampling, as

$$p_i \longrightarrow e^{p_i h} \tag{4.1.9}$$

the stability of $\Delta H_2(z)$ follows from that of $\Delta G_2(s)$. Moreover if $\Delta G_2(s)$ contains only fast dynamics (poles are far away from $j\omega$ axis), $\Delta H_2(z)$ will only have poles close to the origin. In this sense, we may say that fast unmodeled dynamics in continuous time systems are transformed to fast unmodeled dynamics in sampled systems.

For multiplicative uncertainty in continuous time systems

$$\overline{G}(s)= G(s)(1+ \Delta G_1(s)) \tag{3.1.10}$$

the sampled systems are more complicated. It has been suggested that sampled systems are of the form

$$\overline{H}(z)= H(z)(1+ \Delta H_1(z))$$

where $H(z)$ is a nominal model (sampling of G(s)). It would appear that $\Delta H_1(z)$ be stable because of the stability of $\Delta G_1(s)$. This is unfortunately not true.

Fact 2.

Consider the continuous system (4.1.10), then

(1) The sampled system $\overline{H}(z)$ of $\overline{G}(s)$ may be written as

$$\overline{H}(z)= H(z)(1+ \Delta H_1(z)) \tag{4.1.11}$$

where H(z) is the sampling of G(s).

(2) $\Delta H_1(z)$ in (4.1.11) is not necessarily stable. A sufficient condition for $\Delta H_1(z)$ to be stable is that H(z) be minimum phase.

Proof: Write

$$\overline{H}(z)= \frac{n_{\overline{H}}(z)}{d_{\overline{H}}(z)}, \quad H(z)= \frac{n_H(z)}{d_H(z)}$$

then we have

$$\overline{H}(z)= H(z)(1+ \frac{n_{\overline{H}}(z)d_H(z)-d_{\overline{H}}(z)n_H(z)}{d_{\overline{H}}(z)n_H(z)})$$

$$= H(z)(1+ \Delta H_1(z))$$

This completes the first part. Now from the relation (4.1.9) and the assumption that $\Delta G_1(s)$ is stable, we have

$$d_{\overline{H}}(z) = d_H(z)\gamma(z)$$

where $\gamma(z) = (z - e^{p_1 h}) \ldots (z - e^{p_k h})$, a stable polynomial with the p_i's poles of $\Delta G_1(s)$. This implies that

$$\Delta H_1(z) = \frac{n_{\overline{H}}(z) - \gamma(z) n_H(z)}{\gamma(z) n_H(z)} = \frac{n_{\overline{H}}(z)}{\gamma(z) n_H(z)} - 1 \qquad (4.1.12)$$

Then conclusion follows readily.

Q.E.D.

Remarks:

(1) If the relative degree of G(s) is greater than 2 and $(1 + \Delta G_1(s))$ is strictly proper (for example, fast unmodeled dynamics $\frac{1}{1+\epsilon s} = (1 - \frac{\epsilon s}{1+\epsilon s}))$, then $\Delta H_1(z)$ is always unstable for h is small (see Astrom [3]).

(2) Notice that $n_H(z)$ depends only on G(s), but $n_{\overline{H}}(z)$ depends on G(s) and $\Delta G_1(s)$. Consequently, in most cases, cancellations between $n_H(z)$ and $n_{\overline{H}}(z)$ are unlikely. This implies that unstable $n_H(z)$ usually gives rise to unstable $\Delta H_1(z)$.

(3) From this fact, we see that the assumption of stability on multiplicative uncertainty $\Delta H_1(z)$ may not be reasonable for sampled systems and the result of fact (1) may not be applicable to sampled systems.

Now let us focus on the effect of two special but important classes of unmodeled dynamics in continuous systems, fast mode uncertainty and near pole-zero cancellations, on sampled systems. Consider the continuous time systems of the form

$$\overline{G}(s) = G(s)\frac{1}{1+\epsilon s} \qquad \epsilon > 0$$

and

$$\overline{G}(s) = G(s)\frac{s - \alpha + \epsilon}{s - \alpha} \qquad \alpha < 0,\ \epsilon > 0$$

with G(s) being the nominal model.

To begin with, consider the simple system

$$G(s) = \frac{k}{s - p} \qquad (4.1.13)$$

By calculation, its sampled system is

$$H(z) = \frac{k}{p}\,\frac{e^{ph} - 1}{z - e^{ph}} \qquad (4.1.14)$$

The sampled system corresponding to $\overline{G}(s)=\dfrac{k}{(s-p)(1+\epsilon s)}$ is

$$\overline{H}(z)=H(z)\,A(\epsilon)\,\frac{z-B(\epsilon)}{z-e^{-h/\epsilon}} \tag{4.1.15}$$

where H(z) is defined in (4.1.14) and

$$A(\epsilon)=1+\frac{\epsilon p\,(e^{-h/\epsilon}-e^{ph})}{(e^{ph}-1)(1+\epsilon p)} \tag{4.1.16}$$

$$B(\epsilon)=\frac{(e^{ph}-1)e^{-h/\epsilon}+\epsilon p\,e^{ph}(e^{-h/\epsilon}-1)}{(e^{ph}-1)+\epsilon p\,(e^{-h/\epsilon}-1)} \tag{4.1.17}$$

Notice that $A(\epsilon)\to 1$, $B(\epsilon)\to 0$ as $\epsilon\to 0$. By observing this fact and equation (4.1.15), it is seen that fast mode uncertainty in continuous time system becomes near pole-zero cancellation around the origin for the sampled system.

We now extend this to the high order case

$$G(s)=\frac{k(s-z_1)\dots(s-z_m)}{(s-p_1)\dots(s-p_n)} \quad (m<n,\ p_i\neq p_j \text{ for } i\neq j) \tag{4.1.18}$$

$$=\frac{k_1}{s-p_1}+\dots+\frac{k_n}{s-p_n}$$

From equation (4.1.14), we obtain

$$H(z)=\frac{k_1}{p_1}\,\frac{e^{p_1h}-1}{z-e^{p_1h}}+\dots+\frac{k_n}{p_n}\,\frac{e^{p_nh}-1}{z-e^{p_nh}}$$

$$=\frac{n_1(z)+\dots+n_n(z)}{(z-e^{p_1h})\dots(z-e^{p_nh})}=\frac{\alpha_1z^{n-1}+\dots+\alpha_n}{z^n+\beta_1z^{n-1}+\dots+\beta_n} \tag{4.1.19}$$

where $n_i(z)$ is defined as

$$n_i(z)=\frac{k_i}{p_i}\,(e^{p_ih}-1)\prod_{j=1,j\neq i}^{n}(z-e^{p_jh}) \tag{4.1.20}$$

Similar to the first order case, if there is a fast mode uncertainty $\overline{G}(s)=G(s)\,\dfrac{1}{1+\epsilon s}$, we have

$$\overline{H}(z)=\frac{k_1}{p_1}\,\frac{e^{p_1h}-1}{z-e^{p_1h}}\,A_1(\epsilon)\,\frac{z-B_1(\epsilon)}{z-e^{-h/\epsilon}}+\dots+\frac{k_n}{p_n}\,\frac{e^{p_nh}-1}{z-e^{p_nh}}\,A_n(\epsilon)\,\frac{z-B_n(\epsilon)}{z-e^{-h/\epsilon}}$$

$$=\frac{n_1(z)A_1(\epsilon)(z-B_1(\epsilon))+\dots+n_n(z)A_n(\epsilon)(z-B_n(\epsilon))}{(z-e^{p_1h})\dots(z-e^{p_nh})(z-e^{-h/\epsilon})} \tag{4.1.21}$$

For $\epsilon\to 0$, $A_i(\epsilon)\to 1$, $B_i(\epsilon)\to 0$, this implies that there are near pole-zero cancellations between $(z-e^{-h/\epsilon})$ and $(z-B_i)$ (i=1,..n). In summary, we have shown that the fast mode uncertainty $\dfrac{1}{1+\epsilon s}$ in the continuous system becomes a near pole-zero cancellation around the origin in the sampled system.

Now we study the effect of near pole-zero cancellation in continuous systems on the sampled systems. First consider the simple system

$$\overline{G}(s)=\frac{k}{s-p}\,\frac{s-\alpha+\epsilon}{s-\alpha}\qquad \alpha<0$$

its sampled system is

$$\overline{H}(z)=H(z)\;C(\epsilon)\;\frac{z-D(\epsilon)}{z-e^{\alpha h}} \tag{4.1.22}$$

$H(z)$ is defined as before by (4.1.14) and

$$C(\epsilon)=1+\frac{\epsilon}{e^{ph}-1}\left(\frac{1}{\alpha}+\frac{e^{ph}}{p-\alpha}+\frac{pe^{\alpha h}}{\alpha(\alpha-p)}\right) \tag{4.1.23}$$

$$D(\epsilon)=\frac{(e^{ph}-1)e^{\alpha h}+\epsilon\left(\frac{1}{\alpha}(e^{ph}+e^{\alpha h})+\frac{1+e^{\alpha h}}{p-\alpha}e^{ph}+\frac{1+e^{ph}}{\alpha(\alpha-p)}pe^{\alpha h}\right)}{e^{ph}-1+\epsilon\left(\frac{1}{\alpha}+\frac{e^{ph}}{p-\alpha}+\frac{pe^{\alpha h}}{\alpha(\alpha-p)}\right)} \tag{4.1.24}$$

Notice that $C(\epsilon)\rightarrow 1$, $D(\epsilon)\rightarrow e^{\alpha h}$ as $\epsilon\rightarrow 0$, then comparing equations (4.1.14) and (4.1.22), it is seen that the near pole-zero cancellation at α in continuous systems becomes near pole-zero cancellation in sampled systems around $e^{\alpha h}$. Similarly, we extend this to the more general case. Consider

$$\overline{G}(s)=\frac{k(s-z_1)...(s-z_m)}{(s-p_1)...(s-p_n)}\,\frac{s-\alpha+\epsilon}{s-\alpha}\qquad (m<n,\ p_i\neq p_j,\ \text{for } i\neq j) \tag{4.1.25}$$

From equation (4.1.22), we have

$$\overline{H}(z)=\frac{k_1}{p_1}\,\frac{e^{p_1h}-1}{z-e^{p_1h}}\,C_1(\epsilon)\,\frac{z-D_1(\epsilon)}{z-e^{\alpha h}}+...+\frac{k_n}{p_n}\,\frac{e^{p_nh}-1}{z-e^{p_nh}}\,C_n(\epsilon)\,\frac{z-D_n(\epsilon)}{z-e^{\alpha h}}$$

$$=\frac{n_1(z)C_1(\epsilon)(z-D_1(\epsilon))+...+n_n(z)C_n(\epsilon)(z-D_n(\epsilon))}{(z-e^{p_1h})...(z-e^{p_nh})(z-e^{\alpha h})} \tag{4.1.26}$$

where $n_i(z)$'s are defined by (4.1.20). Since $C_i\rightarrow 1$ and $D_i\rightarrow e^{\alpha h}$ as $\epsilon\rightarrow 0$, it can be interpreted as near pole-zero cancellations between $(z-e^{\alpha h})$ and $(z-D_i)$ (i=1,..n). In summary, we have shown that the near pole-zero cancellation at α in the continuous system becomes near pole-zero cancellation around $e^{\alpha h}$ for the sampled system.

Remark:

Notice that in both cases, the effect of unmodeled dynamics is related to the sampling interval h . When h is large, then

$$e^{-h/\epsilon},\ e^{\alpha h}\rightarrow 0$$

i.e. the near pole-zero cancellations will be close to the origin.

We have discussed the uncertainty in sampled systems due to fast mode unmodeled dynamics and near pole-zero cancellations in continuous time systems. We give the interpretations in terms of the poles and zeros of discrete transfer functions. In the following, we will give an explanation using the ARMA model (or equivalently, the state space model).

Consider the continuous system described by (4.1.18)

$$G(s)=\frac{k(s-z_1)\ldots(s-z_m)}{(s-p_1)\ldots(s-p_n)} \quad (m<n,\ p_i\neq p_j,\ \text{for } i\neq j) \tag{4.1.27}$$

The sampled system is, by equation (4.1.19)

$$H(z)=\frac{n_1(z)+\ldots+n_n(z)}{(z-e^{p_1h})\ldots(z-e^{p_nh})}=\frac{\alpha_1 z^{n-1}+\ldots+\alpha_n}{z^n+\beta_1 z^{n-1}+\ldots+\beta_n} \tag{4.1.28}$$

i.e. the ARMA model may be written as

$$\begin{aligned} y(k)= &-\beta_1 y(k-1)-\ldots-\beta_n y(k-n) \\ &+\alpha_1 u(k-1)+\ldots+\alpha_n u(k-n) \end{aligned} \tag{4.1.29}$$

In the presence of fast mode unmodeled dynamics $\frac{1}{1+\epsilon s}$, the sampled system is

$$\overline{H}(z)=\frac{(\alpha_1+o_1(\epsilon))z^n+(\alpha_2+o_2(\epsilon))z^{n-1}+\ldots+(\alpha_n+o_n(\epsilon))z+o_{n+1}(\epsilon)}{z^{n+1}+(\beta_1-e^{-h/\epsilon})z^n+\ldots+(\beta_n-\beta_{n-1}e^{-h/\epsilon})z-\beta_n e^{-h/\epsilon}} \tag{4.1.30}$$

where $o_i(\epsilon)\rightarrow 0$ as $\epsilon\rightarrow 0$.

The equation (4.1.30) implies that the ARMA model becomes

$$\begin{aligned} y(k)= &(e^{-h/\epsilon}-\beta_1)y(k-1)+(\beta_1 e^{-h/\epsilon}-\beta_2)y(k-2)+\ldots+ \\ &(\beta_{n-1}e^{-h/\epsilon}-\beta_n)y(k-n)+\beta_n e^{-h/\epsilon}y(k-n-1)+(\alpha_1+o_1(\epsilon))u(k-1) \\ &+\ldots+(\alpha_n+O_n(\epsilon))u(k-n)+o_{n+1}(\epsilon)u(k-n-1) \end{aligned} \tag{4.1.31}$$

In the presence of near pole-zero cancellation $\frac{s-\alpha+\epsilon}{s-\alpha}$, we have

$$\overline{H}(z)=\frac{(\alpha_1+o(\epsilon))z^n+(\alpha_2-\alpha_1 e^{\alpha h}+o(\epsilon))z^{n-1}+\ldots+(\alpha_n-\alpha_{n-1}e^{\alpha h})z-\alpha_n e^{\alpha h}}{z^{n+1}+(\beta_1-e^{\alpha h})z^n+(\beta_2-\beta_1 e^{\alpha h})z^{n-1}+\ldots+(\beta_n-\beta_{n-1}e^{\alpha h})z-\beta_n e^{\alpha h}} \tag{4.1.32}$$

The ARMA model is given by

$$\begin{aligned} y(k)= &(e^{\alpha h}-\beta_1)y(k-1)+(\beta_1 e^{\alpha h}-\beta_2)y(k-2)+\ldots.+(\beta_{n-1}e^{\alpha h}-\beta_n)y(k-n) \\ &+\beta_n e^{\alpha h}y(k-n-1)+(\alpha_1+o(\epsilon))u(k-1)+\ldots+ \\ &(\alpha_n-\alpha_{n-1}e^{\alpha h}+o(\epsilon))u(k-n)-\alpha_n e^{\alpha h}u(k-n-1) \end{aligned} \tag{4.1.33}$$

where $o(\epsilon)$ represents the high order term in ϵ.

Compare equations (4.1.28) with (4.1.31) and (4.1.33), we have found that the true ARMA model is just a perturbation of the nominal one in the coefficients by $o_i(\epsilon)$ and $\beta_i e^{-h/\epsilon}$ in the case of fast mode unmodeled dynamics and $\beta_i e^{\alpha h}$ and $\alpha_i e^{\alpha h} + o(\epsilon)$ in the case of near pole-zero cancellation case. Also there are extra high order terms in the ARMA model, say, $\beta_n e^{-h/\epsilon} y(k-n-1)$ and $o_{n+1}(\epsilon) u(k-n-1)$ in the case of fast mode unmodeled dynamics and $\beta_n e^{\alpha h} y(k-n-1)$ and $-\alpha_n e^{\alpha h} u(k-n-1)$ in the case of near pole zero cancellation.

4.2 Slow Drift Instability of Adaptive System

Before analyzing slow drift instability of the adaptive system we present instability theorems for one, two and mixed-time scale dynamic systems. The proofs rely on the results in section 2 and an analysis in the continuous time case.

Consider the difference equation of the form

$$x(k+1) = x(k) + \epsilon f(k, x(k), \epsilon) \qquad x(0) = x_0 \in R^n \tag{4.2.1}$$

where $x \in R^n$, $k \in Z_+$, $0 < \epsilon \leq \epsilon_0$ and f is piecewise continuous with respect to time k. For small ϵ, the variation of x with time is slow as compared to the variation of f. Suppose that the averaged system of (4.2.1) exists

$$x_{av}(k+1) = x_{av}(k) + \epsilon f_{av}(x_{av}(k)) \qquad x_{av}(0) = x_0 \tag{4.2.2}$$

then we have the following theorem:

Theorem 4.2.1: (Instability of an Unaveraged One-Time Scale System)

If: the original system (4.2.1) and the averaged system (4.2.2) satisfy assumptions (A1)-(A4) in section 2 along with the additional assumption that there exists a continuously differentiable decrescent function $v(k,x)$ such that

(i) $v(k,0) = 0$

(ii) $v(k,x) > 0$ for some x arbitrarily close to the origin

(iii) $\left\| \frac{\partial v(k,x)}{\partial x} \right\| \leq k_1 \|x\|$ for some $k_1 > 0$

(iv) the difference of $v(k,x)$ along the trajectory (4.2.2) satisfies

$$v(k+1,x) - v(k,x) \geq \epsilon k_2 \|x\|^2 \tag{4.2.3}$$

Then: the unaveraged system (4.2.1) is unstable provided $\epsilon \leq \epsilon_0$ for some $\epsilon_0 > 0$.

The system of the form (4.2.1) studied is to be thought of as a one-time scale system in that the entire state variable x varies slowly in comparison with the rate of variation of f. We also study for the case when only some of the state variables are slowly varying.

Consider the system

$$x(k+1) = x(k) + \epsilon f(k, x(k), y(k), \epsilon) \qquad x(0) = x_0 \tag{4.2.4}$$

$$y(k+1) = A\, y(k) + \epsilon g(k, x(k), y(k), \epsilon) \qquad y(0) = y_0 \tag{4.2.5}$$

where $x \in R^n$, $y \in R^m$ and $A \in R^{m x m}$ as appearing in (2.2.10), (2.2.11) except that A is a constant matrix independent of state variable x. The averaged system is similarly defined as:

$$x_{av}(k+1) = x_{av}(k) + \epsilon f_{av}(x_{av}(k)) \qquad x_{av}(0) = x_0 \tag{4.2.6}$$

Theorem 4.2.2: (Instability of an Unaveraged Two-Time Scale System)

If: the original system (4.2.4), (4.2.5) and its averaged system (4.2.6) satisfy assumptions (B1)-(B6) in section 2 with the additional assumption that there exists a continuously differentiable decrescent function $v(k,x)$ such that

(i) $v(k,0) = 0$

(ii) $v(k,x) > 0$ for some x arbitrarily close to the origin.

(iii) $\left\| \frac{\partial v(k,x)}{\partial x} \right\| \leq k_3 \|x\|$ for some $k_3 > 0$.

(iv) the difference of $v(k,x)$ along the trajectory (4.2.6) satisfies

$$v(k+1,x) - v(k,x) \geq \epsilon k_4 \|x\|^2 \tag{4.2.7}$$

for some $k_4 > 0$.

Then: the unaveraged system (4.2.4), (4.2.5) is unstable provided $\epsilon \leq \epsilon_0$ for some $\epsilon_0 > 0$.

In adaptive systems, the frequently encountered two-time scale system has the following form:

$$x(k+1) = x(k) + \epsilon f(k, x(k), y'(k), \epsilon) \tag{4.2.8}$$

$$y'(k+1) = Ay'(k) + h(k, x(k)) + \epsilon g'(k, x(k), y'(k), \epsilon) \tag{4.2.9}$$

As shown in section 2, the above system can be transformed into the system (4.2.4), (4.2.5) through the use of the coordinate change

$$y(k) = y'(k) - w(k,x) \tag{4.2.10}$$

where $w(k,x)$ is defined to be

$$w(k,x) = \sum_{i=0}^{k-1} A^{k-i-1} h(i,x) \tag{4.2.11}$$

The averaged system of (4.2.8), (4.2.9) will exist if the following limit exists uniformly in s and x, i.e.

$$f_{av}(x) = \lim_{T\to\infty} \frac{1}{T} \sum_{k=s+1}^{s+T} f(k,x,w(k,x),0) \tag{4.2.12}$$

Theorem 4.2.2 is applicable to this case with one more condition (B7) as in section 2.

The proofs of theorem 4.2.1 and 4.2.2 which are omitted here can be obtained through the results in section 2 and [10].

Remark:

The system that appears in (4.2.8), (4.2.9) corresponds to a linearized adaptive system since A is a constant matrix.

To apply the results to the adaptive system, we consider the case where unmodeled dynamics and bounded output disturbances are present. To begin with, we define the equation error $e(k)$

$$e(k) = y(k) - \phi^T(k-d)\theta(k-1) + \eta(k) \tag{4.2.13}$$

where $\eta(k)$ represents the bounded disturbance, as well as the tuned error $e_*(k)$

$$e_*(k) = y(\bar{\theta}_0, k) - y^*(k) \tag{4.2.14}$$

where $y(\bar{\theta}_0, k)$ denotes the tuned output obtained as if the controller parameter $\theta(k)$ were fixed at $\bar{\theta}_0$ (tuned parameter value).

Before we proceed further, we concretize the concept of tuned plant as the leading step in the analysis of stability/instability of adaptive system.

From the discussion of subsection 4.1, we will consider plant with unmodeled dynamics of the form

$$y(k) = \frac{q^{-d}(B(q^{-1}) + \Delta B(q^{-1}))}{(A(q^{-1}) + \Delta A(q^{-1}))} u(k) \tag{4.2.15}$$

where $q^{-d} B(q^{-1}) / A(q^{-1})$ is considered to be nominal plant and $\Delta A(q^{-1})$, $\Delta B(q^{-1})$ are polynomials of q^{-1} with small coefficients. Denote by $H_\theta(q^{-1})$ the closed loop transfer function from $y^*(k+d)$ to $y(\theta, k)$. One can then show that

$$H_\theta(q^{-1}) = \frac{1 + \Delta N_1(q^{-1})}{1 + \Delta N_2(q^{-1}) + \Delta N_3(q^{-1}) + \Delta N_4(q^{-1})} q^{-d} \tag{4.2.16}$$

where ΔN_1, ΔN_2 and ΔN_3 are stable, perturbed rational function of q^{-1}

$$\Delta N_1(q^{-1}) = \frac{\Delta B(q^{-1})}{B(q^{-1})} \tag{4.2.17}$$

$$\Delta N_2(q^{-1}) = \frac{A(q^{-1})\Delta\beta(q^{-1}) + q^{-d}B(q^{-1})\Delta\alpha(q^{-1})}{B(q^{-1})} \tag{4.2.18}$$

$$\Delta N_3(q^{-1}) = \frac{\Delta A(q^{-1})\beta(q^{-1}) + q^{-d}\Delta B(q^{-1})\alpha(q^{-1})}{B(q^{-1})} \tag{4.2.19}$$

$$\Delta N_4(q^{-1}) = \frac{\Delta A(q^{-1})\Delta\beta(q^{-1}) + q^{-d}\Delta B(q^{-1})\Delta\alpha(q^{-1})}{B(q^{-1})} \tag{4.2.20}$$

with $\alpha(q^{-1})$ and $\beta(q^{-1})$ are as defined in (3.2.5), (3.2.6) corresponding to the nominal parameter θ_0, whereas $\Delta\alpha(q^{-1})$ and $\Delta\beta(q^{-1})$ are perturbed polynomials corresponding to the frozen parameter θ.

By the small gain theorem, the transfer function $H_\theta(q^{-1})$ is stable provided the controller parameter θ (fixed) and plant uncertainties $\Delta A(q^{-1})$ and $\Delta B(q^{-1})$ are such that

$$\sup_{\omega \in [-\pi,\pi]} |\Delta N_i(e^{j\omega})| < \frac{1}{3} \qquad i = 2,3,4 \tag{4.2.21}$$

Thus, the collection of such θ constitutes a robust parameter set as denoted by $\Pi(\theta_0)$ (the continuous counterpart appears in [10,19]).

As a result, the tuned error $e_*(k)$ defined in (4.2.14) can be evaluated through the difference of transfer functions, i.e.

$$H_\theta(q^{-1}) - q^{-d} = q^{-d}E(q^{-1}) \tag{4.2.22}$$

where

$$E(q^{-1}) = \frac{\Delta N_1(q^{-1}) - (\Delta N_2(q^{-1}) + \Delta N_3(q^{-1}) + \Delta N_4(q^{-1}))}{1 + \Delta N_2(q^{-1}) + \Delta N_3(q^{-1}) + \Delta N_4(q^{-1})} \tag{4.2.23}$$

Also

$$\int_{-\pi}^{\pi} S_{e_*}(d\omega) = \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 S_{y^*}(d\omega) \tag{4.2.24}$$

where $S_{e_*}(d\omega)$ denotes the spectral density function of $e_*(k)$.

As illustrated in the previous discussion in subsection 4.1, the unmodeled error of a sampled-data system $|\Delta A(e^{j\omega})|$ and $|\Delta B(e^{j\omega})|$ tend to be smaller as the sampling period gets to be larger in both the case of fast mode unmodeled dynamics and near pole-zero cancellation (provided the unmodeled pole in the 1st case is high enough and the pole & zero in the 2nd case are close enough).

To relate the input frequency with the aforementioned unmodeled error $|\Delta A(e^{j\omega})|$ and $|\Delta B(e^{j\omega})|$, we consider only the case of fast mode unmodeled dynamics and refer to the following diagram

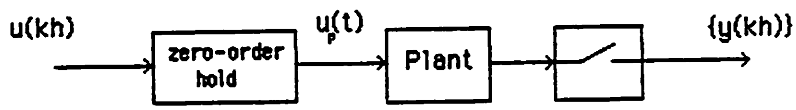

Fig. 3

Suppose the input $u(t)$ is a pure sinusoidal signal, say, $sin(\omega_0 t)$ with $\omega_0 \leq \frac{1}{2}\omega_s$ where ω_s is the sampling frequency which will be assumed large enough in our case. By sampling theory, the immediate input to the plant, $u_p(t)$, which is also the output of the zero-order-hold device, contains spectral lines at frequencies $\omega_0 + m\omega_s$, however, with magnitudes decaying as m tends to be large. In particular, if $\omega_0/\omega_s << 1$, then $u_p(t)$ will have the spectral line at frequency ω_0 as the only significant frequency content. This implies that $y_p(t)$ will be closer to the nominal plant output and, in turn, that $|\Delta A(e^{j\omega})|$ and $|\Delta B(e^{j\omega})|$ are smaller as ω_0 decreases.

Due to these facts, given $\rho > 0$, there exist a sampling frequency and a reference trajectory signal with appropriate frequency support such that

$$\int_{-\pi}^{\pi} S_{e_\theta}(d\omega) \leq \rho \int_{-\pi}^{\pi} S_{y^*}(d\omega) \tag{4.2.25}$$

and with $\theta \in \Pi(\theta_0)$ (i.e. $H_\theta(q^{-1})$ remains stable) where $e_\theta(k) = y(\theta,k) - y^*(k)$. The collection of such θ, corresponding to such a sampling frequency and reference trajectory signal, will then be called tuned parameter set as defined in [10,19].

Let $\overline{\theta}_0$ be chosen, such that (4.2.25) is satisfied, as a tuned value, then $H_{\overline{\theta}_0}$ will be defined as the tuned plant transfer function. We rewrite the equation error $e(k)$ in (4.2.13) as:

$$\begin{aligned} e(k) = &\{ y(k) - y(\overline{\theta}_0, k)\} + \{ e_*(k) + \eta(k)\} \\ &+ \{ y^*(k) - \phi^T(k-d)\,\theta(k-1)\} \end{aligned} \tag{4.2.26}$$

As shown in [1], the 1st term can be reexpressed as:

$$y(k)-y(\bar{\theta}_0,k) = -\{ H_{\bar{\theta}_0}(q^{-1})\,q^d \}\, \phi^T(k-d)\,\tilde{\theta}(k-d) \tag{4.2.27}$$

where $\tilde{\theta}(k) = \theta(k)-\bar{\theta}_0$ and the 3rd term

$$y^*(k)-\phi^T(k-d)\,\theta(k-1) = \phi^T(k-d)\,(\,\theta(k-d)-\theta(k-1)\,) \tag{4.2.28}$$

, by which equation (4.2.26) becomes

$$\begin{aligned} e(k) = & -\{ H_{\bar{\theta}_0}(q^{-1})\,q^d \}\, \phi^T(k-d)\,\tilde{\theta}(k-1) + (\, e_*(k) + \eta(k)\,) \\ & + \{\, 1 - H_{\bar{\theta}_0}(q^{-1})\,q^d \}\, \phi^T(k-d)\,(\,\theta(k-d)-\theta(k-1)\,) \end{aligned} \tag{4.2.29}$$

As a result of (4.2.29), parameter adaptation of the projection type becomes

$$\begin{aligned} \tilde{\theta}(k) = \; & \tilde{\theta}(k-1) - \epsilon \frac{\phi(k-d)\,\{ H_{\bar{\theta}_0}(q^{-1})\,q^d \}\,(\,\phi^T(k-d)\,\tilde{\theta}(k-1)\,)}{1+\epsilon\,\phi(k-d)^T\,\phi(k-d)} \\ & + \epsilon \frac{\phi(k-d)\,(\, e_*(k) + \eta(k)\,)}{1+\epsilon\,\phi^T(k-d)\,\phi(k-d)} \\ & + \epsilon \frac{\phi(k-d)\,\{\, 1 - H_{\bar{\theta}_0}\,q^d \}\,(\,\phi^T(k-d)\,(\,\tilde{\theta}(k-d)-\tilde{\theta}(k-1)\,)}{1+\epsilon\,\phi^T(k-d)\,\phi(k-d)} \end{aligned} \tag{4.2.30}$$

Notably, the analysis of local stability/instability of system (4.2.30) can be confined to just studying the 1st term on the R.H.S. of eq. (4.2.30) of ϵ order provided $e_*(k)$, $\eta(k)$ are sufficiently small so that the 2nd term can be discarded, and provided ϵ is sufficiently small such that the last term of ϵ^2 order can be neglected.

In order to make the system (4.2.30) ready for averaging analysis, we formulate the error model analogous to (3.2.18), (3.2.19) except for the aforementioned truncation and linearization, i.e.

$$\tilde{x}(k+1) = A(\bar{\theta}_0)\,\tilde{x}(k) + \frac{\partial h(k,\theta)}{\partial \theta}\Big|_{\theta=0} \cdot \tilde{\theta}(k) \tag{4.2.31}$$

$$\tilde{\theta}(k) = \tilde{\theta}(k-1) - \epsilon \frac{\phi_{\bar{\theta}_0}(k-d)\,\{ H_{\bar{\theta}_0}(q^{-1})\,q^d \}\,(\,\phi^T_{\bar{\theta}_0}(k-d)\,\tilde{\theta}(k-1)\,)}{1+\epsilon\,\phi^T_{\bar{\theta}_0}(k-d)\,\phi_{\bar{\theta}_0}(k-d)} \tag{4.2.32}$$

where $\phi_{\bar{\theta}_0}$ is similarly defined in (3.2.27). By averaging results in section 2, the averaged system of (4.2.31) is

$$\tilde{\theta}_{av}(k) = (I - \epsilon\, R_{\bar{\theta}_0}(0)\,)\,\tilde{\theta}_{av}(k-1) \tag{4.2.33}$$

where

$$R_{\bar{\theta}_0}(0) = \lim_{T\to\infty} \frac{1}{T} \sum_{i=s+1}^{s+T} \left\{ \phi_{\bar{\theta}_0}(i-d)\,\{ H_{\bar{\theta}_0}(q^{-1})\,q^d \}\,(\,\phi^T_{\bar{\theta}_0}(i-d)\,) \right\}$$

$$s \geq d-1 \tag{4.2.34}$$

If $R_{\bar{\theta}_0}(0)$ has eigenvalues with negative real part, then the simplest Lyapunov function that satisfies the additional assumptions (i)-(iii) has the form

$$v(\tilde{\theta}_{av}) = \tilde{\theta}_{av}^T P \tilde{\theta}_{av} \tag{4.2.35}$$

where P is the matrix containing at least one positive eigenvalue , satisfying

$$(I - \epsilon R_{\bar{\theta}_0}(0))^T P (I - \epsilon R_{\bar{\theta}_0}(0)) - P = Q \tag{4.2.36}$$

for some positive definite Q. Applying Theorem 4.2.2, we immediately conclude that linearized, truncated adaptive system (4.2.30), (4.2.31) and hence the original adaptive system are unstable.

Remark:

In fact, the averaged system (4.2.33) is exponentially stable if $H_{\bar{\theta}_0}(q^{-1})\, q^d$ is strictly positive real (discrete) and $\phi_{\bar{\theta}_0}(k)$ is persistently exciting for sufficiently small ϵ, which implies that the original adaptive system is BIBO stable around the tuned parameter $\bar{\theta}_0$.

Suppose that the reference trajectory $y^*(k)$ is stationary with finite number of spectral lines, then the frequency domain expression of (4.2.34) assumes the form

$$R_{\bar{\theta}_0}(0) = \frac{1}{2\pi} \sum_{i=1}^{m_0} n_{\bar{\theta}_0}(e^{-j\omega_i}) H_{\bar{\theta}_0}(e^{j\omega_i})\, e^{-j\omega_i d}\, n_{\bar{\theta}_0}(e^{j\omega_i})\; r_i^2$$

$$+ \frac{1}{2\pi} \sum_{i=1}^{m_1} \bar{n}_{\bar{\theta}_0}(e^{-j\omega_i}) H_{\bar{\theta}_0}(e^{j\omega_i})\, e^{-j\omega_i d}\, \bar{n}_{\bar{\theta}_0}(e^{j\omega_i})\; \eta_i^2 \tag{4.2.37}$$

where $n_{\bar{\theta}_0}(q^{-1})$ and $\bar{n}_{\bar{\theta}_0}(q^{-1})$ are column transfer functions from $y^*(k)$ and $\eta(k)$ to $\phi_{\bar{\theta}_0}(k)$ respectively provided $y^*(k)$ and $\eta(k)$ don't contain common spectral lines.

Remark:

To comply with the old notation in (3.2.27), we have

$$n_{\bar{\theta}_0}(q^{-1}) = \frac{C(q^{-1})}{C^T(q^{-1})\cdot\bar{\theta}_0} \quad \text{and} \quad \bar{n}_{\bar{\theta}_0}(q^{-1}) = \frac{\bar{C}(q^{-1})}{C^T(q^{-1})\cdot\bar{\theta}_0} \tag{4.2.38}$$

where

$$C^T(q^{-1}) = (q^{-d}B(q^{-1}), \cdots, q^{-(n+d-1)}B(q^{-1}), A(q^{-1}), \cdots, q^{-(m+d-1)}A(q^{-1})) \tag{4.2.39}$$

and

$$\overline{C}(q^{-1}) = (0, \ldots, 0, \alpha_{\overline{\theta}_0}(q^{-1})A(q^{-1}), \cdots, \alpha_{\overline{\theta}_0}(q^{-1})q^{-(m+d-1)}A(q^{-1})) \qquad (4.2.40)$$

where $\alpha_{\overline{\theta}_0}(q^{-1})$ is as defined in (3.2.5) with α_i, $i = 0,1,\ldots,n-1$ corresponding to the tuned parameter $\overline{\theta}_0$, which result in

$$R_{\overline{\theta}_0}(0) = \frac{1}{2\pi}\sum_{i=1}^{m_0} \frac{C(e^{-j\omega_i})\, C^T(e^{j\omega_i})}{|C^T(e^{-j\omega_i})\cdot\overline{\theta}_0|^2} H_{\overline{\theta}_0}(e^{j\omega_i})\, e^{-j\omega_i d}\, r_i^2$$

$$+ \frac{1}{2\pi}\sum_{i=1}^{m_1} \frac{\overline{C}(e^{-j\omega_i})\, \overline{C}^T(e^{j\omega_i})}{|C^T(e^{-j\omega_i})\cdot\overline{\theta}_0|^2} H_{\overline{\theta}_0}(e^{j\omega_i})\, e^{-j\omega_i d}\, \eta_i^2 \qquad (4.2.41)$$

Before we state a theorem which relates the instability behavior with the frequency content of the reference trajectory signal, we make the following definitions.

Definition 4.2.3: Good Signals, Bad Signals

A stationary signal is said to be good signal if its spectral support $\subset \left\{\omega \,|\, -90^o < \angle H_{\overline{\theta}_0}(e^{j\omega})\, e^{j\omega d} < 90^o \right\}$. A stationary signal is called bad if the spectral support $\subset \left\{\omega \,|\, \angle H_{\overline{\theta}_0}(e^{j\omega})e^{j\omega d} < -90^o \text{ or } \angle H_{\overline{\theta}_0}(e^{j\omega})e^{j\omega d} > 90^o \right\}$

Theorem 4.2.4:

Suppose the linearized, truncated system described by (4.2.31) and (4.2.32) is not persistently excited by good signals, then a bad signal with sufficiently small or large magnitude, will result in the instability of the adaptive system.

Proof: The proof of the theorem follows from the fact that

$$\lambda(I - \epsilon R_{\overline{\theta}_0}(0)) = 1 - \epsilon\lambda(R_{\overline{\theta}_0}(0)) \qquad (4.2.42)$$

where λ represents eigenvalue, and the results in [10].

Q.E.D.

Incidentally, the least squares type adaptation algorithm also possess similar instability property. The following corollary may be easily proven by the theorems similar to those of section of [10].

Corollary 4.2.5:

If all conditions in Theorem 4.2.4 are satisfied with the adaptive law changed to be least squares type with forgetting factor plus the total spectral lines due to either input or output disturbance are more than $n+m+d$, then a bad signal with sufficiently small magnitude will result in the instability of the adaptive system.

To illustrate the slow drift phenomena we discussed above, consider an example due to Rohrs, namely a continuous 1st order nominal plant along with a pair of fast unmodeled modes as follows

$$\overline{G}(s) = \frac{2}{s+1}\,\frac{229}{s^2+30s+229}$$

The nominal and true sampled systems (with h =0.5) are given by $H(z)$ and $\overline{H}(z)$ respectively.

$$H(z) = \frac{0.787}{z-0.607}$$

$$\overline{H}(z) = \frac{0.612(z^2+0.286z+3.9\cdot10^{-4})}{(z-0.607)(z^2-5.977\cdot10^{-4}z+3.059\cdot10^{-7})}$$

We choose the tuned parameter value $\overline{\theta}_0$ to be $(0.61, 0.80)^T$. Here, we examine the slow drift instability in the case where $y^*(k)=0$ and output disturbance $d(k)=sin(wk)$. It is not hard to show that the conditions of the theorem 4.2.2 are satisfied when $w=\pi$ due to the fact that $d(k)$ is a bad signal by definition (4.2.3). The slow drift of parameters is verified by the simulation as shown in Fig.4.

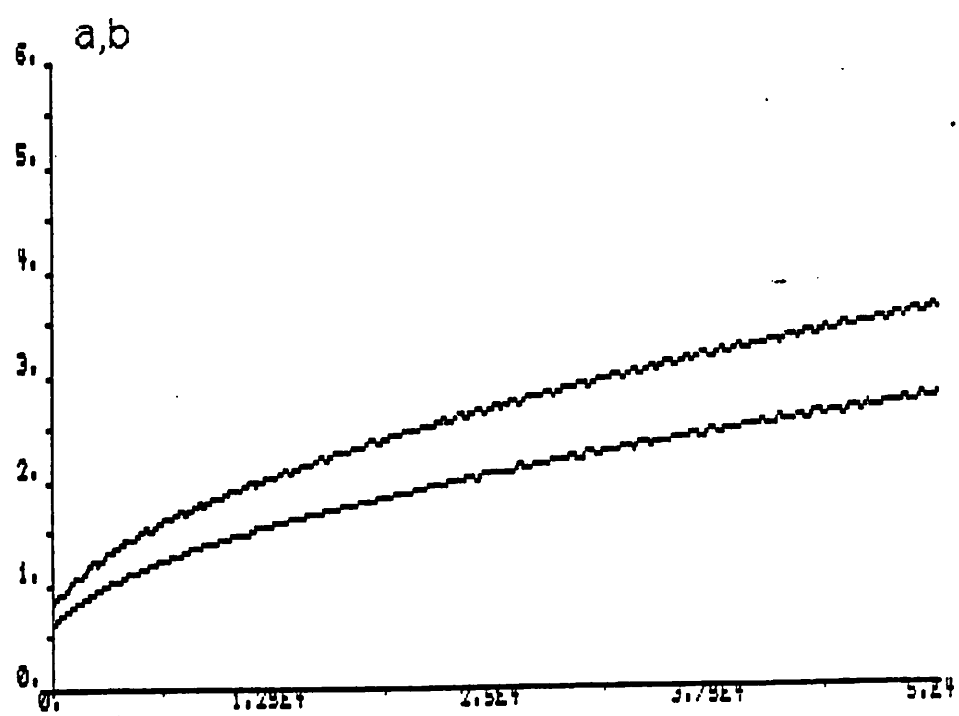

Fig. 4 Parameter Drifting ($\theta=(a,b)^T$)

5. Concluding Remarks:

There were two important parts to our paper: (1) The complete extension of averaging results from the continuous time to the discrete time situation. (2) An initiation of the discussion of the effect of sampling on unmodeled dynamics in a continuous time plant and the effects of these on the overall robustness of the adaptive system. This second part, we feel, is the start of a larger program on the systematic study of the adaptive control of sampled data systems, which is initiated in Bai & Sastry [4], Goodwin et al [11] and Praly et al [25].

6 References

(1) B.D.O. Anderson, R.R.Bitmead, C.R.Johnson,Jr., P.V.Kokotovic, R.L.Kosut, I.M.Y.Mareels, L.Praly, and B.D.Riedle 'Stability of Adaptive System: Passivity and Averaging Analysis' 1986 The MIT Press Cambridge,MA

(2) K. J. Astrom, "Interactions Between Excitation and Unmodeled Dynamics," IEEE Trans. on AC, Volume AC-30, No. 9, September 1985, pp 889-891

(3) K.J.Astrom, P.Hagander and J.Sternby "Zeros of Sampled Systems' Automatica Vol.20.No.1 1981 pp31-38

(4) E.W. Bai and S.S.Sastry 'Adaptive Stabilization of Sampled Systems' Mem. No. UCB/ERL M86/52 University of California, Berkeley

(5) E.W. Bai and S.S.Sastry 'Persistency of Excitation, Sufficient Richness and Parameter Convergence in Discrete Time Adaptive Control' Systems and Control Letters Vol.6 1985 pp153-163

(6) Marc Bodson, S. Sastry, B.D.O. Anderson, I. Mareels and R. Bitmead, "Nonlinear Averaging Theorems, and the Determination of Parameter Convergence Rates in Adaptive Control," System and Control Letters 7 (1986), pp 145-157

(7) S. Boyd and S. Sastry, "Necessary and Sufficient Conditions for Parameter Convergence in Adaptive Control," Proc. of the Ames-Berkeley Conf. on Non-Linear Dynamics and Control, Math-Sci Press, Brookline, Massachusetts, 1984, pp 81-101

(8) J.C. Doyle and G.Stein 'Multivariable Feedback Design: Concept for a Classical Modern Synthesis' IEEE Trans. on Ac Vol.Ac-25 1981 pp4-6

(9) L.C. Fu, M. Bodson and S. Sastry, "New Stability Theorems for Averaging and Their Application to Convergence Analysis of Adaptive Identification and Control Schemes," Proc. 24th IEEE Conf. on Decision and Control, Ft. Lauderdale, FL, 1985, pp 473-477

(10) L.C.Fu and S.S.Sastry 'Slow Drift Instability in Model Reference Adaptive System--an Averaging Analysis' to appear in Int. J. of Contr.

(11) G.C.Goodwin, R.L.Leal,D.Q.Mayne and R.H.Middleton 'Rapproachment Between Continuous and Discrete Model Reference Adaptive Control' Automatica Vol.22 No.2 pp1-9

(12) G.C.Goodwin,P.J.Ramadge and P.E.Caines 'Discrete Time Multivariable Adaptive Control' IEEE Trans. on AC Vol. AC-25 1980 pp449-456

(13) J. K. Hale, "Ordinary Differential Equations," Kreiger, Molaban (Florida), 1980

(14) D.J.Hill and P.J.Moylan 'Connections Between Finite-Gain and Asymptotic Stability' IEEE Trans. on AC Vol. AC-25 1980 pp931-936

(15) P. Ioannou and K. Tsakalis 'A Robust Discrete Time Adaptive Controller' submitted to 25th Conf. on Decision and Control

(16) P.V.Kokotovic,B.D.Riedle and L.Praly 'On a Stability Criterion for Slow Adaptation' Systems and Contr. Letters Vol.6 pp7-14 1985

(17) R.Kosut 'Method of Averaging for Adaptive Systems' Research paper

(18) R. Kosut, B. D. O. Anderson and I.M.Y.Mareels "Stability Theory for Adaptive Systems: Method of Averaging and Persistency of Excitation," IEEE Trans. on AC, to appear

(19) R. Kosut and C.R. Johnson 'An Input-output View of Robustness in Adaptive Control' IFAC 1984, pp569-581

(20) G.Kreisselmeier and B.D.O. Anderson 'Robust Model Reference Adaptive Control' IEEE Trans. on AC Vol.AC-31 1986 pp127-133

(21) W.E.Larimore, S.Mahmood and R.K.Mehra 'Adaptive Model Algorithmic Control ' Proc. IFAC Workshop on Adaptive Systems in Control and Signal Processing, June 1983 San Francisco

(22) L.Praly 'Reduction Principle in Slowly Adaptation Discrete Time Systems' Research paper

(23) L.Praly'A geometric Approach for The Local Analysis of a One Step Ahead Adaptive Controller' Proc. 4th Yale Workshop on Applications of Adaptive System Theory, Yale University,CT

(24) L.Praly 'Local Stability/Instability of Slowly Adapted System in Presence of Periodic Inputs' Research paper

(25) L.Praly, S.T.Hung and D.S.Rhode 'Towards a Direct Scheme for a Discrete -Time Control of a Minimum Phase Continuous Time System' Proc. of 24th Conf. on Decision and Control Ft.Lauderdale 1985 pp1188-1191

(26) L.Praly and D.Rhode 'A Local Analysis of a One Step Ahead Adaptive Controller' Proc.24th Conf.on Deci. and Contr. 1985 Ft.Lauderdale, pp1862-1867

(27) B.D. Riedle 'Integral Manifolds of Slow Adaptation' Ph.D. Thesis 1986 University of Illinois,Urbana-Champaign

(28) B.D.Riedle 'Stability Bounds for Slow Adaptation:an Integral Approach' Proc. of 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing, Lund, Sweden,1986 pp61-66

(29) B. D. Riedle and P. V. Kokotovic, "Stability Analysis of An Adaptive System with Unmodeled Dynamics," Int. J. Control, Vol.41 No.2 1985, pp.389-402

(30) B. D. Riedle and P.V. Kokotovic, 'A Stability-Instability Boundary for Disturbance-Free Slow Adaptation and Unmodeled Dynamics' IEEE Trans. on AC. Vol.AC-30 No. 10 1985 pp

1027-1030

(31) B.D. Riedle and P.V. Kokotovic 'Integral Manifold Approach to Slow Adaptation' Proc.of 24th Conf. on Dec. and Contr. Ft. Lauderdale,FL Dec.1985 pp147-151

(32) M.Vidyasagar and A.Vannelli 'New Relationships Between Input-Output and Lyapunov Stability' IEEE Trans. on AC Vol.AC-27 1982 pp481-483